**Research Papers :**

* Particle Swarm Optimization-Based Training Algorithm for Feedforward Neural Networks (2009) by Nianjun Liu and Zhenxing Wang. Journal of Software, vol. 4, no. 10, pp. 1076-1083, 2009. [Link](https://www.jsoftware.us/vol4/jsw0410-1266.pdf)
* A Genetic Algorithm Based Neural Network Ensemble for Time Series Prediction (2008) by Shuangyan Liu, Lei Wang, and Yanping Hu. Expert Systems with Applications, vol. 35, no. 3, pp. 1542-1547, 2008. [Link](https://www.sciencedirect.com/science/article/pii/S0957417407004143)
* Ant Colony Optimization Based Neural Network for Short-term Load Forecasting (2007) by D. K. Chaturvedi and V. Mukherjee. International Journal of Electrical Power & Energy Systems, vol. 29, no. 6, pp. 445-452, 2007. [Link](https://www.sciencedirect.com/science/article/pii/S0142061506003381)
* Hybrid Ant Colony Optimization Algorithm for Neural Network Training (2006) by Lin-Yu Tseng and Gwo-Hshiung Tzeng. Expert Systems with Applications, vol. 31, no. 2, pp. 378-386, 2006. [Link](https://www.sciencedirect.com/science/article/pii/S0957417405004414)
* Genetic Algorithms and Particle Swarm Optimization for Feature Selection in Classification of Large Datasets (2006) by J. K. Aggarwal and S. K. Singh. Pattern Recognition Letters, vol. 27, no. 8, pp. 860-874, 2006. [Link](https://www.sciencedirect.com/science/article/pii/S0167865505003907)

**Performance Comparison**

From the implementation of the neural network and the three optimization algorithms for the given dataset plus the research done to implement them, I have arrived at the following result

1. Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) are effective in finding a global optima and thus may be more suitable for feature selection and parameter optimization in neural networks. Among those from what I have read genetic algorithm is a bit slow to converge [for large datasets] and thus PSO is a better choice in terms of them time but doesn’t guarantee global optima at the end .
2. Ant Colony Optimization (ACO) is the algorithm that showed the best results among the three, mainly because of the discrete search space [research papers].
3. And I really didn’t read enough about CA and thus wasn’t able to implement it to my fullest ability.